

DATA ANALYTICS REFERENCE DOCUMENT	
Document Title:	Descriptive Statistics
Document No.:	1540829111
Author(s):	Rita Raher
Contributor(s):	

REVISION HISTORY

Revision	Details of Modification(s)	Reason for modification	Date	By
0	Draft release	Document description here	2018/10/29 16:05	Rita Raher

# Descriptive Statistics

Descriptive statistics describes the data. It is distinguished from inferential statistics, in that descriptive statistics aims to summarize a sample, rather than use the data to learn about the population that the sample of data is thought to represent. This generally means that descriptive statistics, unlike inferential statistics, is not developed on the basis of probability theory, and are frequently nonparametric statistics.

Some measures that are commonly used to describe a data set are measures of central tendency and measures of variability or dispersion. Measures of central tendency include the mean, median and mode, while measures of variability include the standard deviation (or variance), the minimum and maximum values of the variables, kurtosis and skewness.

Methods for visually presenting summary statistics include tables, charts, and graphical plots.

## Descriptive statistics using Pandas

```
import pandas as pd
df.describe()
```

## Central Tendency

In statistics, a central tendency (or measure of central tendency) is a central or typical value for a probability distribution. It may also be called a center or location of the distribution.

The most common measures of central tendency are:

- **mean** - the average
- **median** - the value separating the higher half from the lower half of a data sample
- **mode** - The set of data values is the value that appears most often

```
import numpy as np
# finding the mean
np.mean(df)
```

Last update:  
2020/06/20  
14:39

help:statistics:descriptive-statistics

<http://hdip-data-analytics.com/help/statistics/descriptive-statistics>

```
# mean of a column
np.mean(df['col'])

# finding the median
np.median(df)
```

## Graphical Plots

Graphical plots are interesting in that they pictorially convey a large amount of information in a concise way that allows for quick interpretation and understanding of the data.

The Box Plot

From:

<http://hdip-data-analytics.com/> - **HDip Data Analytics**

Permanent link:

<http://hdip-data-analytics.com/help/statistics/descriptive-statistics>

Last update: **2020/06/20 14:39**